Conformational perturbation of SARS-CoV-2 Spike protein using N-acetyl cysteine, a molecular scissor: A probable strategy to combat COVID-19

Utsab Debnath[1], Varun Dewaker[2], Yenamandra S. Prabhakar[2], Parthasarathi Bhattacharyya[3], Amit Kumar Mandal[4]

[1] School of Pharmaceutical Technology, Adamas University, Kolkata, West Bengal, India, Pin–700126

[2] Medicinal and Process Chemistry Division, CSIR-Central Drug Research Institute, Lucknow, India, Pin–226031

[3] Institute of Pulmocare & Research, New Town, Kolkata, West Bengal, India, Pin – 7000156

[4] Department of Biological Science, Indian Institute of Science Education and Research Kolkata, Nadia, West Bengal, India – 741246

*Corresponding Author:
Dr. Amit Kumar Mandal
Associate Professor,
Department of Biological Sciences,
Indian Institute of Science Education and Research Kolkata
Mohanpur, Nadia, West Bengal, India-741246
E-mail: amitkm@iiserkol.ac.in

Abstract:

The infection caused by Severe Acute Respiratory Syndrome–CoronaVirus-2 (SARS-CoV-2) resulted in a pandemic across the globe with a huge death toll. The symptoms from SARS-CoV-2 appear somewhat similar to the SARS-CoV-1 infection that appeared in early 21st century but the infectivity is far higher for the SARS-CoV-2. The virus attaches itself to exposed human epithelial cells through the spike protein. Recently discovered crystal structure of the complex of spike protein of SARS-CoV-2 with human angiotensin-converting enzyme 2 (ACE2) receptor indicated that the virus binds with the host cell very strongly. We hypothesized that the perturbation of the functionally active conformation of spike protein through the reduction of a solvent accessible disulfide bond (Cys391-Cys525) that provides its structural architecture, may

be a feasible strategy to disintegrate the spike protein from ACE2 receptor and thereby prevent the infection. Using *in silico* platform we showed that N-acetyl cysteine (NAC), a drug used as antioxidant and mucolytic agent, binds in the close proximity of above disulfide bond. The reduction of the disulfide bond via thiol/disulfide exchange, followed by covalent conjugation of NAC perturbed the stereo specific orientations of interacting key residues of spike protein. This resulted in threefold weakening in the binding affinity of spike protein with ACE2 receptor. This opens avenues for exploring the effect of NAC *in vitro, ex vivo and in vivo* and on successful observation of the similar effect as *in silico*, the intervention of NAC may be translated in the pharmacoprevention and treatment of Corona virus disease 2019.

Key words: SARS-CoV-2, Spike protein, Disulfide bonds, N-Acetyl Cysteine (NAC), Molecular Docking, MD simulation, Angiotensin-converting enzyme 2 (ACE2)

1. Introduction:

Corona virus disease 2019 (COVID-19) is caused by the highly pathogenic RNA virus, SARS-CoV-2.[1] Similar to SARS-CoV-1 and Middle East Respiratory Syndrome-CoronaVirus (MERS-CoV), the infection caused by SARS-CoV-2 shows a range of symptoms such as dry cough, fever, headache, dyspnoea and pneumonia. However, the rapid rate of infection caused by SARS-CoV-2 with a mortality rate of 3-5% has resulted in a pandemic across the globe.[2] Bat coronavirus RaTG13 has been found to be the closest relative of the SARS-CoV-2.[2] Phylogenetic analyses of the coronavirus genomes showed that SARS-CoV-2 belongs to the Beta coronavirus genus. The genome of SARS-CoV is a single-stranded RNA consisting of about 30 kb nucleotides.[3] SARS-CoV-2 encodes for four major structural proteins, namely spike protein (S), membrane protein (M), envelope protein (E), and nucleocapsid protein (N).

Coronavirus infection starts with the binding of viral particles to the host cells. The entry of the SARS-CoV-2 into human cells initiates through the transmembrane spike protein that forms trimers protruding from the virus cell surface.[3, 4] SARS-CoV-2 uses receptor-binding domain (RBD) of spike protein to bind strongly with ACE2 to enter into the host cells with a dissociation equilibrium constant, $K_d \approx 15$ nM.[1, 4] ACE2 is a type I membrane protein expressed in multiple organs such as lungs, heart, kidneys, and intestine.[1] Its primary physiological role is in the

maturation of angiotensin (Ang), a peptide hormone that controls vasoconstriction and blood pressure.[1] Down regulation of ACE2 has been reported to be associated with cardiovascular diseases.[1]

Thus, the receptor recognition of the spike protein is critical to the development of COVID-19. Crystal structure of the spike protein and ACE2 complex showed that there are thirteen hydrogen bonds and two salt bridges at SARS-CoV-2 RBD-ACE2 interface.[2] The stereo specific orientations of the interacting amino acid residues of spike protein are provided by its structural architecture, where four intra-molecular disulfide bonds between eight cysteine residues contribute. Among these four disulfide linkages, three are present in the N-terminal domain (NTD) of the protein (Cys336–Cys361, Cys379–Cys432 and Cys391–Cys525), which stabilize the β sheet structure and the fourth disulfide bond (Cys480–Cys488) joins the loops at the distal end of the receptor binding motif that belongs to the C-terminal domain (CTD) of the spike protein.[2]

We hypothesized that the functionally active conformation of spike protein could be perturbed and rendered inactive through the reduction of its solvent accessible disulfide bond. Thus, based on a thiol/disulfide exchange mechanism between the cystine residue of the spike protein and NAC, a commonly used drug as an antioxidant and a mucolytic agent, it might be possible to reduce the solvent accessible disulfide bond between Cys391 and Cys525, which might lead to conformational change in the RBD of spike protein.[5] We carried out a cysteine residue based molecular docking study with NAC in the vicinity of the most accessible disulfide bond (Cys391-Cys525) at NTD of spike protein (PDB ID: 6LZG). Subsequently, Molecular Dynamics (MD) simulation was performed up to 100 ns to analyze the structural stability of spike protein-NAC complex. Upon covalent conjugation of NAC with Cys525, a comparative *in silico* analysis was performed to predict the stereo specific orientation of interacting residues of RBD of spike protein with ACE2 in the NAC conjugate of spike protein.

2. Materials and Methods:

2.1. *Solvent accessibility analysis based on cysteine residues*

To analyze the solvent accessibility of cysteine residues connected via disulfide bonds, spike protein was extracted from the crystal structure of spike protein-ACE2 receptor (human)

complex (PDB ID: 6LZG).[6] After selecting the protein, VEGAZZ based pocket analysis was carried out specifically for eight cysteine residues which are available in the spike protein using default parameters.[7]

2.2. *Pocket identification and Molecular Docking*

To identify the probable binding pockets of ligand free spike protein, CASTp web server was used by considering solvent probe sphere 1.4 Å. The pocket identification analysis is based on recent theoretical and algorithmic results of Computational Geometry, which includes delaunay triangulation, alpha shape, and discrete flow.[8] After selecting the pocket site, the docking study was performed in AUTODOCK 4.2.[9] Prior to the docking, N-acetyl cysteine (NAC) was generated in Sybyl-X 1.3.[10] Subsequently NAC was subjected to optimization for its geometry using the Powell energy minimization algorithm, Gasteiger-Huckel charges, and 0.001 kcal/(mol. Å) as convergence criteria.[11] The protein coordinates of spike protein from the spike protein– ACE2 receptor (human) co-crystal (PDB code: 6LZG) was considered for investigating the binding mode of NAC. For docking study, the protein was prepared in SYBYL by making use of the same procedure as adopted in case of study molecule. Here, residues (Cys480, Cys488 and their surrounding residues) based docking method was carried out to dock the NAC molecules near to the ACE2 receptor binding domain. The prepared spike protein and the corresponding study molecule were considered in AUTODOCK for the docking experiments. The grid size for the search of docking space was set at 60 X 60 X 60 Å^3 distributed around the binding domain with a default grid spacing of 0.375 Å.[12] The Lamarckian genetic algorithm was used for docking the molecules. The docking was done with 100 runs. The molecules were allowed to flexibly dock into the protein coordinates (6LZG) to take their final conformation. The best docked conformers were collected from a population of 150 samples from 2.5 million energy evaluations. Along with AUTODOCK, Pymol and MOE software were used to visualize the interactions of protein and docked molecules.[12, 13]

2.3. *Molecular Dynamics (MD) simulations.*

In NAMD2.7 MD simulations were carried out on docking model to study the stability of these systems and dynamics of molecular interactions between spike protein and NAC.[14] To do this, the protonation state of histidine residues of the protein at physiological pH (7.4) were assigned

using the H++ web server.[15] CHARMM22 force Field (with CMAP correction) was used to parameterize the protein.[16] This force field yielded improved dynamical and structural properties of proteins in MD simulations. Parameters of the ligands were generated via the ParamChem web server.[17] The MD simulations were carried out on under periodic boundary conditions using the solvated (water box) protein. The prepared protein-ligand system was run for 100 ns using multi-core and CUDA supported NAMD package. The NPT technique (Pressure 1 atm; Temperature 310 K) with 12 Å as cut-off for non-bonded atom interactions, Particle Mesh Ewald algorithm for long-range electrostatic forces and Langevin's piston in 'on' condition were used to run the MD simulation.[18] The computation was executed with the standard parameter files (CHARMM22) of MD simulation, default arguments and an internal dielectric constant value 1.6.[19] The time step for dynamics integration was set to 2 fs. The simulation trajectories were recorded at 1000 steps (2 ps) intervals. The system was adequately minimized for 70 ps followed by subjected to equilibration run for 100 ns. Visualization and evaluation of the trajectories were done in VMD 1.9.3.[20] Making use of standard tcl scripts the RMSD and Rg of the respective protein trajectories were analyzed.

2.4. *Protein-protein docking:*

Protein-protein docking study was carried via HDOCK server.[21] This exercise was performed for crystal structure of ligand free spike protein with ACE2 receptor complex as well as MD simulated NAC docked spike protein with ACE2 receptor complex to understand the difference in their binding affinities towards ACE2 receptor. This study was based on a hybrid algorithm of template-based modeling and ab initio free docking method. Prodigy server was used to determine ΔG and Kd values of the predictive protein-protein interaction contacts.[22] Pymol software was used to visualize the interactions between two proteins. Furthermore, NAC molecule was manually conjugated with free thiol group of Cys525 residue by reducing the most solvent accessible disulfide bond (Cys391-Cys525) to form a new Cys525-S-S-NAC linkage. This was executed on the basis of distance matrix analysis in between NAC and cysteine residues after MD simulation study. The experiment was carried out to check the difference in binding affinities with ACE2 receptor through additional protein-protein docking simulation.

3. Result and Discussion:

An attempt to impair the binding of the spike protein of SARS-CoV-2 with human ACE2 receptor might be a feasible strategy to inhibit Cornonavirus infection. In the present study, we addressed this problem through the perturbation of the functionally active conformation of spike protein by reducing a disulfide bridge that provides the structural architecture of spike protein. We observed that reduction of a disulfide bridge located at a distant position from the site of interaction between spike protein and ACE2 receptor, followed by the covalent conjugation of N-acetyl cysteine with the Cys525 resulted in a significant perturbation of the conformation of spike protein that might lead to inhibition in its binding with ACE2 receptor.

Spike protein interacts with ACE2 receptor through its C-terminal domain (CTD). The main interacting residues are following: Leu 455, Tyr473, Tyr489, Gln493, Asn501, Tyr505 (Figure 1). Crystal structure of spike protein (PDB ID: 6LZG) showed that out of nine cysteine residues, eight cysteines are covalently connected by four disulfide bridges (Cys336–Cys361, Cys379–Cys432, Cys391–Cys525 and Cys480–Cys488). The disulfide bond between Cys480 and Cys488 resides in the CTD of spike proteins and the remaining three disulfide bonds belong to NTD of the protein. The crystal structure of the spike protein-ACE2 complex showed that Cys480-Cys488 disulfide bond of the spike protein is located closer to the interacting residues of ACE2 receptor in the complex, whereas other disulfide bonds are at a distant from the site of interaction in the complex. Table-1 shows that the solvent accessible surface area (SASA) of different cysteine residues in the crystal structure (PDB ID: 6LZG). Analysis of SASA indicated that among all four disulfide bonds in the spike protein, the bond between Cys391 and Cys525 is the most accessible one to the solvent molecules in solution.

One of the strategies to reduce the disulfide bond in a protein molecule is through the thiol/disulfide exchange mechanism executed via free thiol group (-SH) of a reducing agent such as reduced glutathione (GSH, a natural antioxidant present *in vivo*), N-acetyl cysteine (an antioxidant and mucolytic agent) etc.[5, 23] GSH is a tripeptide and it is larger in size compared to NAC. In fact, NAC is being prescribed as a drug in last couple of decades to increase the GSH level within cells as it provides one of the constituent amino acid, cysteine, of GSH through its penetration across the cell membranes followed by the deacetylation with the enzyme

6

deacetylase.[24] The main focus of the present study was to explore the mechanism of action of NAC through the docking of NAC within spike protein followed by the reduction of a disulfide bond in its vicinity that eventually might lead to unfolding of spike protein. To perform this experiment spike protein was extracted from its complex with ACE2 receptor (PDB ID: 6LZG) and the four aforementioned disulphide bonds of spike protein were identified.

Before docking study of NAC, SASA and pocket identification analysis were carried out to find the most acceptable binding site of NAC molecule within the spike protein. It was observed that high SASA were present near two following disulfide bond forming cysteine residues: one near the vicinity of RBD site (Cys480, 17.29 Å) and other one in NTD (Cys391, 57 Å) (Table 1). Moreover, pocket site identification study showed five probable binding sites with diverse pocket volume as well as pocket surface areas (Table 2, Figure 2). Comparative pocket volume area analysis clearly indicated that the pocket solvent accessible area near to the RBD site was less (74.09 Å^2) than that of N-terminal site (89.68 Å^2) (Table 2, Figure 2). However, during molecular docking study, second preferable pocket was considered for NAC binding as it was closer to the RBD site. The distance between the centroids of RBD and the second preferable site was ~ 15 Å whereas the most preferable site was relatively far away from the RBD site with a distance of ~ 40 Å. To run the docking study, cysteine and other nearby pocket residues-based grid was generated. The docking study indicated that NAC occupy a position in the binding pocket with considerable binding affinity (ΔG = -7.32 Kcal.mol^{-1}). The docking pose indicated a strong H-bond interaction between NAC and two following residues, Pro479 and Asn481 into the binding pocket (Table 3, Figure 3). However, due to lack of any aromatic group in the NAC moiety, there was no π-π stacking interaction with any aromatic amino acid residue present at the binding site. Only polar and hydrophobic interactions were observed between NAC and nearby residues including Phe486, Asn481, Cys480 and Pro479 (Table 3). The distance between the centroid of NAC and Cys480 was 6.0 Å whereas the distance from the residue Cys488 was 4.2 Å. Subsequently, to check the stability of NAC-spike protein complex, MD simulation study was performed for 100 ns. MD simulation study showed that the trajectories of protein and NAC were not very much close to each other due to shifting of NAC from docking site to N-terminal site of the spike protein. RGs confirmed that these systems followed a harmonious swirl throughout the dynamics (Figures are shown in supporting information). After MD simulation, the coordinates

of most important ACE2 interacting amino acids were changed (Figure 4). Interestingly, NAC became unglued from the docked site after 3 ns run of MD simulation and travelled across the solvent system. Finally, it was bound to the N-terminal site of spike protein. The distance between initial and final positions of NAC was 59.2 Å. The new location of NAC was close to the disulfide bond in between Cys391 and Cys525 residues (Table 3, Figure 3). The binding affinity of NAC with spike protein was improved by two H-bond interactions with the pocket residue Thr523. In addition, it showed hydrophobic interactions with other pocket residues such as Val362 and Ala522. From the last trajectory of the MD simulation study, the distance between centroid of NAC with Cys391 and Cys525 were observed to be 4.4 Å and 3.2 Å respectively (Figure 5). Thus, it improved the probability of thiol/disulfide exchange between free thiol group (–SH) of NAC and Cys391-Cys525 disulfide bond.

Predicted dissociation equilibrium constant, K_d values of protein-protein interaction clearly revealed that upon NAC conjugation with Cys525, the binding affinity of spike protein with ACE2 was decreased by three folds with a concomitant decrease in the release of free energy (ΔG) by 5 Kcal.mol^{-1} (Table 4). The change in the number of polar-polar as well as polar-apolar/non polar interactions at the interface between two proteins clearly indicated the molecular insights of the differential binding affinities associated with NAC conjugation (Table 4). MD simulation study did not have any constrain to break the disulfide bond and to form a new disulfide bond with a ligand. Thus, the reduction of Cys391-Cys525 bond and subsequent conjugation between NAC with Cys525, the closest residue, was made manually by maintaining the protein coordinates (Figure 6). Upon reduction of the disulfide bond, the spike protein was unfolded to a significant extent. However no similar type of unfolding event was observed for other three disulfide bonds of spike protein (Supporting information). Thus, it clearly indicated that the Cys391-Cys525 disulfide bond significantly contributes in the structural integrity of spike protein. In order to inhibit the binding of spike protein with ACE2 receptor, the above information provided important inputs on the cleavage of the disulfide bond between Cys391 and Cys525 residues by the action of NAC. After NAC conjugation with Cys525, protein minimization was carried out. The conformational change of the spike protein indicated that the unfolding of the spike protein led to increase in the distance between C-terminus and N-terminus of the spike protein from 9.9 Å, as per crystal structure of the complex, to 130.2 Å (Figure 6).

This resulted in a significant perturbation of the stereo specific orientation of the interacting residues in the RBD of spike protein with those of ACE2 receptor which eventually manifested in a decreased binding affinity of spike protein with ACE2 receptor (Table 5, Figure 6).

4. Conclusion:

In the present study, docking analysis of NAC followed by MD simulation indicated that the binding of NAC with spike protein is thermodynamically favoured in the pocket that is in close proximity to the most accessible disulfide bonds between Cys391 and Cys525 residues of spike protein. Upon reduction of this disulfide bond through the thiol disulfide exchange mechanism with free sulfhydryl group of NAC followed by the covalent conjugation of NAC with Cys525 resulted in a perturbation of the functionally active conformation of spike protein where the spatial orientations of the participating key residues in binding of spike protein with ACE2 receptor were significantly distorted. This led to three-fold reduction in the binding affinity of the spike protein with ACE2 receptor, which might be translated in the inhibition in infection caused by the SARS-CoV-2. Thus, the aforementioned action of NAC might be explored sequentially *in vitro*, *ex vivo* and *in vivo* and on successful observation as *in silico*, NAC may be used as a drug in the inhibition of the infection caused by SARS-COV-2 in COVID 19.

6. Supporting Information: Supporting information include (1) Distance comparison between C and N-terminal of the spike protein, before and after NAC conjugation. (2) Positional difference in the interaction points of selected important amino acids of spike protein with ACE2 receptor. (3) Structural deformations after individual thiol formation at four different disulfide positions (across eight cysteine residues) of spike protein (4) Radius of Gyration of spike protein after MD simulation (5) Radius of Gyration of spike-ligand (NAC) after MD simulation and (6) Root mean square fluctuation of ligand (NAC) during MD simulation.

7. Supporting Videos: Supporting videos include (1) Removal of NAC from the docking site of the viral spike protein and (2) binding affinity of NAC with NTD of the viral spike protein during MD simulation up to 100 ns.

8. Notes: Authors declared no conflict of interest.

9. References:

(1) Renhong, Y.; Yuanyuan, Z.; Li, Y.; Lu, X.; Yingying G.; Qiang, Z. Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science*, **2020**, 367: 1444–1448.

(2) Jun, L.; Jiwan, G.; Jinfang, Y.; Sisi, S.; Huan, Z.; Shilong, Fan.; Qi, Z.; Xuanling, S.; Qisheng, W.; Linqi, Z.; Xinquan, W. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature*, **2020**, 581, 215–220.

(3) Junwen, L.; Yue, L.; Xiaolu, J.; Leiliang, Z. Spike protein recognition of mammalian ACE2 predicts the host range and an optimized ACE2 for SARS-CoV-2 infection. *Biochem Biophys Res Commun*. **2020**, 21, 526(1), 165–169.

(4) Alexandra C, W.; Young-Jun Park, M.; Alejandra, T.; Abigail, W.; Andrew, T.; McGuire, D. V. Structure, Function, and Antigenicity of the SARSCoV-2 Spike Glycoprotein. *Cell*, **2020**, 180, 1–12.

(5) Aldinia, G.; Altomare, A.; Baron, G.; Vistoli, G.; Carini, M.; Borsani, L.; Sergio, F. N-Acetylcysteine as an antioxidant and disulphide breaking agent: the reasons why? *Free Radical Research*. **2018**, 52, 751–762.

(6) Wang, Q.; Zhang, Y.; Wu, L.; Niu, S.; Song, C.; Zhang, Z.; Lu, G.; Qiao, C.; Hu, Y.; Yuen, K.Y.; Wang, Q.; Zhou, H.; Yan, J.; Qi, J. Structural and Functional Basis of SARS-CoV-2 Entry by Using Human ACE2. *Cell*, **2020**, 181, 894-904.e9.

(7) Pedretti A.; Villa L.; Vistoli G. VEGA - An open platform to develop chemo-bio-informatics applications, using plug-in architecture and script programming. *J Comput Aided Mol Des.*, **2004**, 18, 167-173.

(8) Binkowski, T. A.; Naghibzadeh, S.; Liang, J. CASTp: Computed Atlas of Surface Topography of Proteins. *Nucleic Acids Res*. **2003**, 31, 3352−3355.

(9) Morris, G.M.; Huey R, Lindstrom, W.; Sanner, M.F.; Belew, R.K.; Goodsell, D.S.; Olson, A.J. AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J Comput Chem*. **2009**, 30, 2785–2789.

(10) Sybyl-X 1.3, St. Louis, 2010. http://www.tripos.com.

(11) Debnath, U.; Kumar, P.; Agarwal, A.; Kesharwani, A.; Gupta, S.K.; Katti, S.B. N-hydroxy-substituted 2-aryl acetamide analogs: A novel class of HIV-1 integrase inhibitors. *Chem Biol Drug Des*. **2017**, 90, 527–534.

(12) Seeliger, D.; de Groot, B.L. Ligand docking and binding site analysis with PyMOL and Autodock/Vina. *J Comput Aided Mol Des.* **2010**, 24, 417–422.

(13) MOE; The Molecular Operating Environment from Chemical Computing Group Inc., Montreal, Quebec, Canada; http://www.chemcomp.com.

(14) Phillips, J. C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R. D.; Kale, L.; Schulten, K. Scalable Molecular Dynamics with NAMD. *J. Comput. Chem.* **2005**, 26, 1781−1802.

(15) Gordon, J. C.; Myers, J. B.; Folta, T.; Shoja, V.; Heath, L. S.; & Onufriev, A. H++: A server for estimating pKas and adding missing hydrogens to macromolecules. *Nucleic Acids Res.* **2005**, 33(Web Server), W368–371.

(16) Brooks, B. R.; Bruccoleri, R. E.; Olafson, B. D.; States, D. J.; Swaminathan, S.; & Karplus, M. CHARMM: A program for macromolecular energy, minimization, and dynamics calculations. *J. Comput. Chem.*, **1983**, 4(2), 187–217.

 (17) https://www.paramchem.org/

(18) Dewaker, V.; Srivastava, P.N.; Verma, S.; Prabhakar, Y.S. Molecular dynamics study of HDAC8-largazole analogues co-crystals for designing potential anticancer compounds. *J Biomol Struct Dyn.* **2019**, 38, 1197-1213.

(19) Li, Y.; Cong, Y.; Feng, G.; Zhong, S.; Zhang, J. Z. H.; Sun, H.; Duan, L. The impact of interior dielectric constant and entropic change on HIV-1 complex binding free energy prediction. *Structural Dynamics* (Melville, N.Y.), **2018**, 5(6), 064101.

(20) Humphrey, W.; Dalke, A.; Schulten, K. VMD: Visual molecular dynamics. *J Mol Graph Model.* **1996**, 14(1), 33–38.

(21) Yan, Y.; Zhang, D.; Zhou, P.; Li, B.; Huang S.Y. HDOCK: a web server for protein-protein and protein -DNA/RNA docking based on a hybrid strategy. *Nucleic Acids Res.* **2017**, 45, W365-373,.

(22) Xue, L. C.; Rodrigues, J. P.; Kastritis, P. L.; Bonvin, A. M.; Vangone, A. PRODIGY: a web server for predicting the binding affinity of protein-protein complexes. *Bioinformatics*, **2016**, 32(23), 3676-3678.

(23) Srivastava, D; Sarma, GRK; Kukkuta, D'Souza, D; Muralidharan, M; Srinivasan, K; Mandal, AK; Characterization of residue-specific glutathionylation of CSF proteins in multiple sclerosis – A MS-based approach; *Analytical Biochemistry*, **2019**, 564-565, 108-115

(24) Bürger, M; Chory1, J; Structural and chemical biology of deacetylases for carbohydrates, proteins, small molecules and histones. *Commun Biol.* **2018**, 217, 1-11.

Legend of Figures:

Figure 1: (**A**) Representing the position of RBD site at S1 domain of SARS-CoV-2 . (**B**) Selected important binding sites of protein-protein interaction in between ACE2 receptor (green color) and viral spike protein (cyan color). NTD indicates N-terminal domain where three disulphide bonds (-S-S-) are located within six cysteine residues (Cys336-Cys361; Cys379-Cys432; Cys391-Cys525). The other disulphide bond (-S-S-, Cys480-Cys488) is placed near to ACE2 receptor binding site.

Figure 2: Predictive binding pockets of spike protein. Different colour indicates the position of binding pockets along with pocket volume and surface area for sovent accessbility.

Figure 3: Figures of different binding sites of NAC before and after MD simulation. (**A**) and (**B**) represents the docked poses of NAC before MD simulation. (**A**) indicates 3D view of docked NAC along with their binding residues near to ACE2 receptor binding domain. Red dotted line (---) shows strong H-bond interaction of NAC with Cys 480. (**B**) shows 2D view of docked NAC along with surrounding residues. Green colour shows van der Waal interaction sites whereas violet colour pointed out the polar interaction sites. Proximity contour on each atom of NAC molecule indicates the ligand exposure sites. Similarly figures (**C**) and (**D**) display the binding site of NAC after MD simulation. Figure (**E**) RMSD is the plot analysis of protein (blue) and ligand (brown). Figure (**F**) pointed out the positional distance (59.2Å) between centroids of NAC molecules before (cyan) and after (green) MD simulation.

Figure 4: Figure (**A**) indicates the superimposition of MD simulated spike protein (green in colour) with crystal form (cyan in colour) (RMSD: 1.58Å), whereas figure (**B**) shows the corresponding positional and rotational changes of six important amino acid residues located in the ACE2 recptor binding sites of spike protein.

Figure 5: Distance (in Å) of NAC with disulphide bond forming cysteine residues (Cys 391 and Cys 525) at N-terminal site of the spike protein after MD simulation.

Figure 6: Difference in protein–protein interaction between ACE2 receptor and spike protein before and after NAC conjugation via thiol exchange mechanism: (A) shows interaction between the crystal forms of ACE2 receptor and NAC free spike protein whereas figure (B) shows *in silico*

protein–protein interactions between ACE2 receptor and NAC conjugated spike protein. The red dotted lines (---) in (A) and (B) displays the diatance between C terminal and N terminal of spike protein, before (9 Å) and after (130.2Å) NAC conjugation.
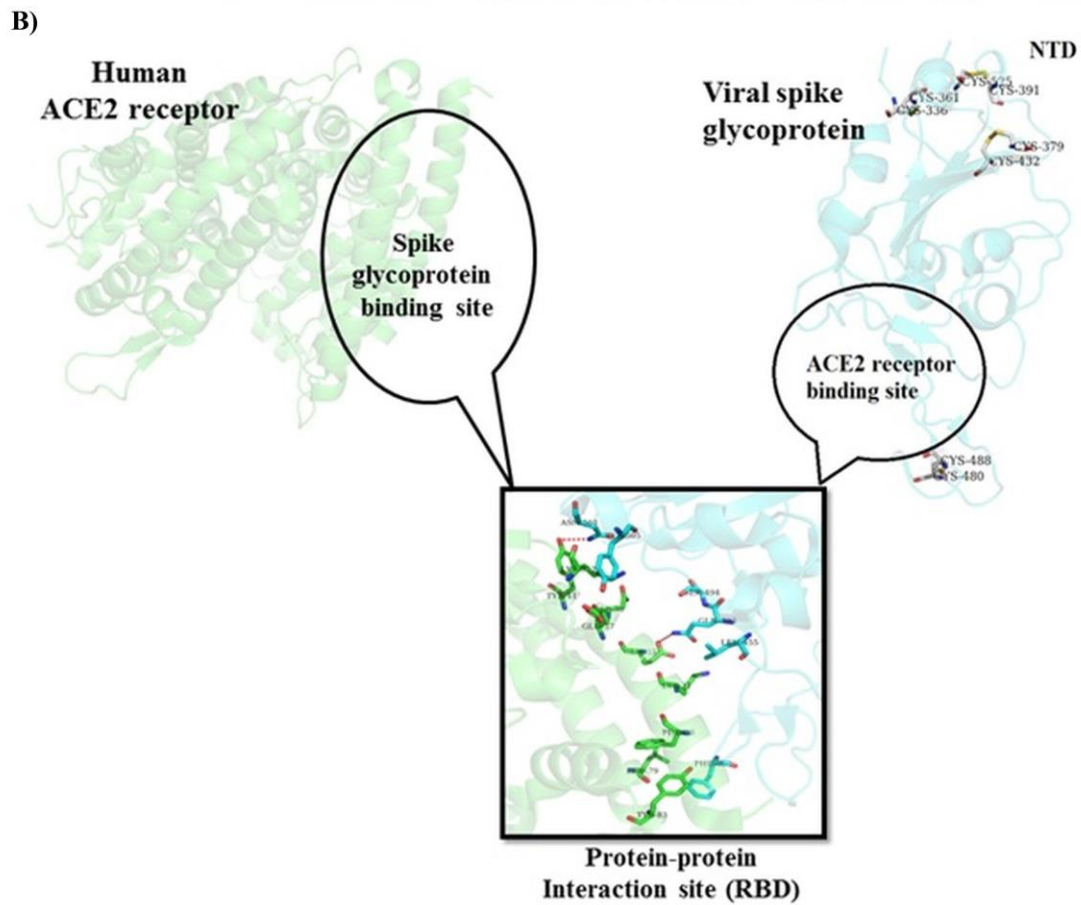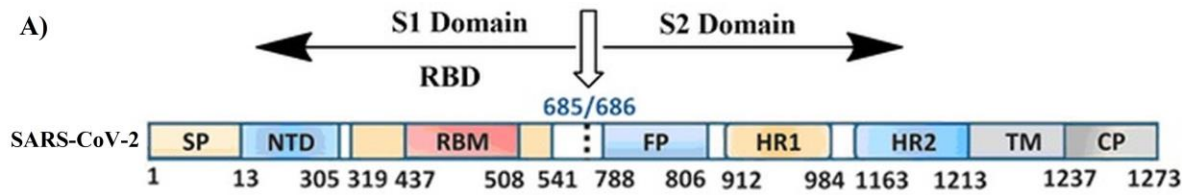
Legend of Tables:

Table 1: Cysteine residues based solvent accessibility analysis of spike protein using a crystal structure (PDB ID: 6LZG)
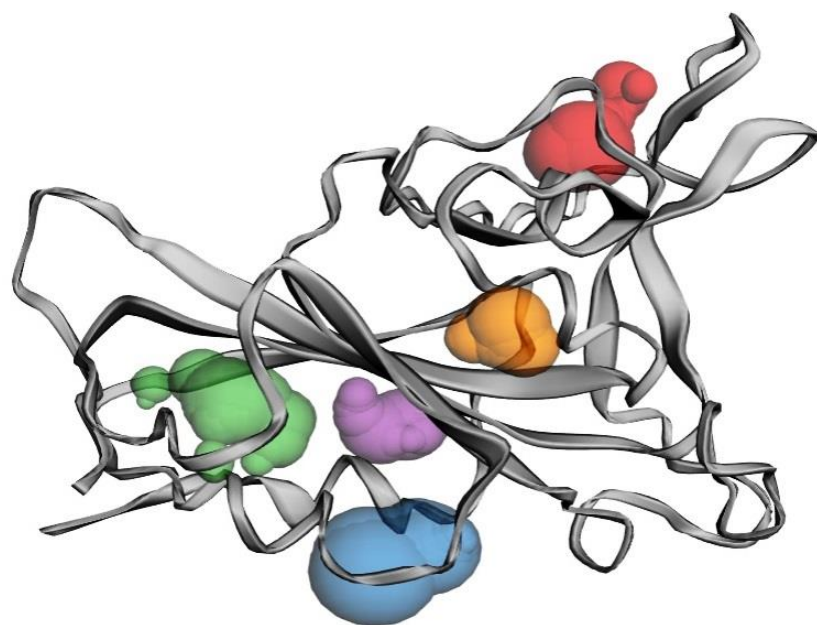
Table 2: Binding pocket identification and analysis

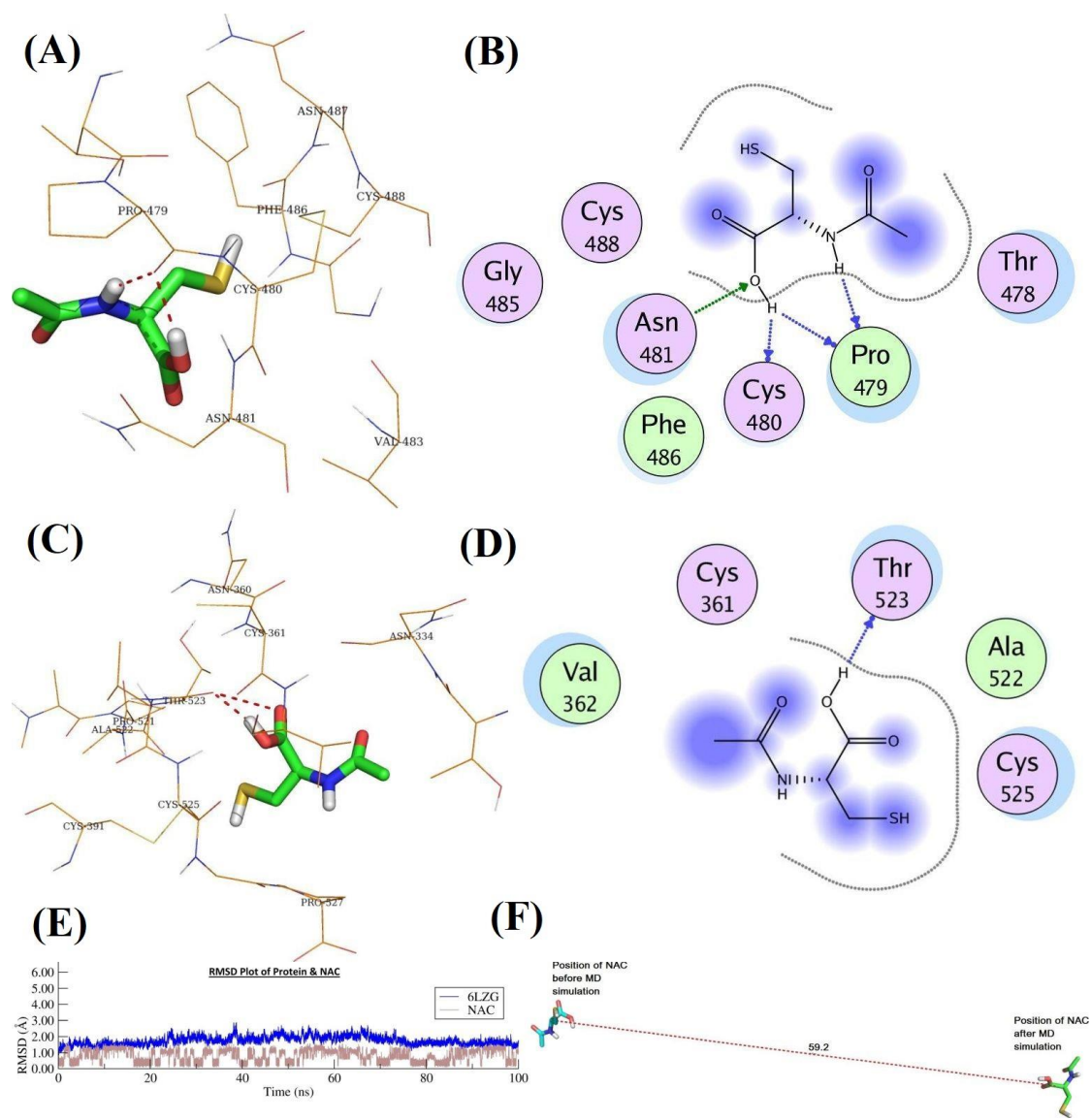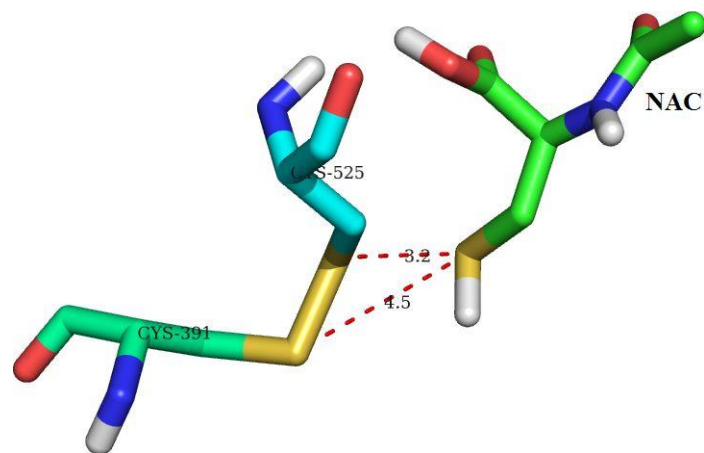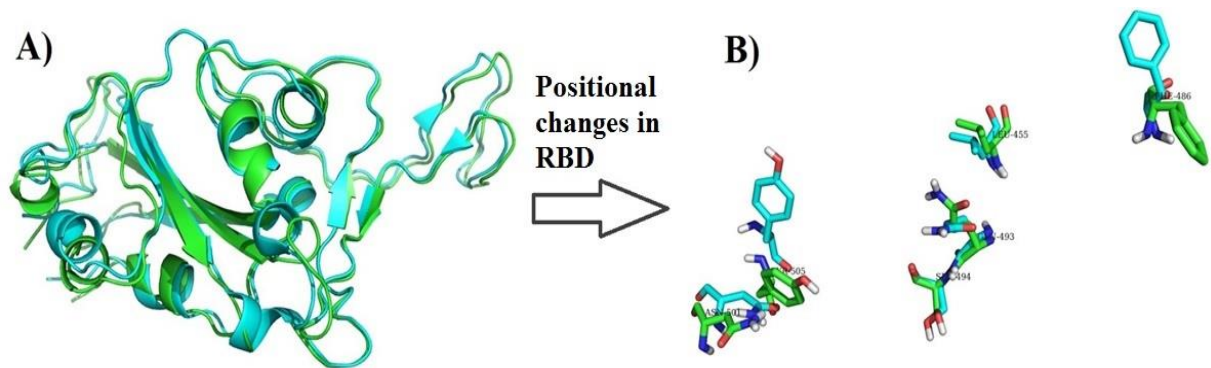**Table 3:** NAC binding pocket analysis

Table 4: Details of protein-protein interaction contacts (ICs)

Table 5: Predictive interaction sites after disulfide bond formation with NAC

A)

B)

Protein-protein
Interaction site (RBD)

(A)

ASN-487
CYS-488
PHE-486
PRO-479
CYS-480
ASN-481
VAL-483

(B)

HS
Gly 485
Cys 488
Thr 478
Asn 481
Phe 486
Cys 480
Pro 479

(C)

ASN-360
CYS-361
ASN-334
THR-523
PRO-521
ALA-522
CYS-525
CYS-391
PRO-527

(D)

Cys 361
Thr 523
Ala 522
Val 362
Cys 525
SH

(E)

RMSD Plot of Protein & NAC

6.00
5.00
4.00
3.00
2.00
1.00
0.00

RMSD (Å)

0   20   40   60   80   100
Time (ns)

6LZG
NAC

(F)

Position of NAC before MD simulation

59.2

Position of NAC after MD simulation

16

A)

B)

Positional changes in RBD

NAC

CYS-525

CYS-391

3.2

4.5

ACE2 receptor

Spike protein

TYR-473
TYR-489
LEU-455
GLN-493
SER-494
TYR-505
ASN-501

NT   CT

Cysteine 525   Cysteine 391

NAC

H_2O

$2H^{\oplus}$

NAC conjugated
Cysteine 525

Cysteine 391

After disulphide (-S-S-)
bond formation in between
NAC and Cys 525 via thiol
exchange mechanism

ACE2 receptor

NT

NAC congugated
with Cys 525

Spike protein

130.2

TYR-505
ASN-501

SER-494
GLN-493

LEU-455

GLU-471   TYR-489
PHE-486

CT

18

Table 1:

Residue-position	ASA(Å^2)^{a,b}	Δ^sG (Kcal.mol^{-1})^c
Cys336	12.30	-0.13
Cys361	11.10	0.05
Cys379	25.53	-0.28
Cys391	57.12	1.55
Cys432	6.61	0.27
Cys480	17.29	0.05
Cys488	7.91	0.27
Cys525	44.74	1.27

[a]ASA = Column ASA indicates the solvent-accessible surface area, in Å^2, for the respective residues.

[b]Å^2 = A square angstrom (Å^2) is a non-SI (non-System International) measurement unit of area with sides equal to one angstrom (1Å).

[c]Δ^sG = Column Δ^sG indicates free energy release associated with the hydrophobic interaction between individual residues and solvent, in kcal/mol.

Note: BSA (buried surface area) is zero for all residue

Table 2:

Pocket_ID	Pocket color	Solvent-accessible Pocket volume area (Å^2)	Solvent-accessible Pocket surface area (Å^2)	Pocket residues
Pocket_1	Red	32.521	74.091	Arg454, Phe456, Arg457, Lys458, Asp467, Ser469, Glu471, Ile472, Tyr473, Gln474, Pro491.
Pocket_2	Blue	30.015	40.774	Phe342, Asn343, Ser371, Ser373, Phe374, Trp436.
Pocket_3	Green	27.141	89.686	Tyr365, Ser366, Leu368, Tyr369, Phe377, Cys379, Val382, Ser383, Pro384, Lys386, Leu387, Asn388, Leu390, Phe392, Cys432, Leu513, Phe515, Val524, Cys525.
Pocket_4	Purple	13.347	35.351	Glu340, Val341, Ala344, Arg346, Phe347, Ala348, Asn354, Lys356, Ala397, Ser399.
Pocket_5	Orange	9.777	28.094	Thr376, Lys378, Tyr380, Val407, Arg408, Ile410, Ala411, Val433.

Table 3:

NAC binding pocket study after docking			NAC binding pocket study after MD simulation		
Pocket residues	H-bond interaction sites	Hydrophobic interaction sites	Pocket residues	H-bond interaction sites	Hydrophobic interaction sites
Cys488, Gly485, Thr478, Phe486, Pro479, Cys480, Asn481, Val483	Pro479, Cys480, Asn481	Phe486, Pro479	Asn334, Val362, Cys361, Cys391, Thr523, Ala522, Pro527, Cys525	Thr523	Val362, Ala522

Table 4:

Spike glycoprotein-ACE2 receptor(human) interaction before NAC binding (crystal form)		
Sl. No.		
1	ΔG (kcal.mol^{-1})	-12.5
2	K_d (M) at 25.0 °C	7.2E-10
3	Number of ICs (charged-charged)	3
4	Number of ICs (charged-polar)	10
5	Number of ICs (charged-apolar)	20
6	Number of ICs (polar-polar)	5
7	Number of ICs (polar-apolar)	23
8	Number of ICs (apolar-apolar)	9

Spike glycoprotein-ACE2 receptor(human) interaction after NAC conjugation (*in silico* form)		
Sl. No.		
1	ΔG (kcal mol^{-1})	-7.48
2	K_d (M) at 25.0 °C	2.2E-09
3	Number of ICs (charged-charged)	3
4	Number of ICs (charged-polar)	7
5	Number of ICs (charged-apolar)	15
6	Number of ICs (polar-polar)	4
7	Number of ICs (polar-apolar)	14
8	Number of ICs (apolar-apolar)	6

Table 5:

Interacting residues of ACE2 receptor	Interacting residues of spike protein
Lys114, Lys131, Gln493, Pro499, Gly504, Gln506, Asn394	Gly482, Phe464, Asp67, Asp206, Ser44, Asn394

Table of Content